wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:10 PM BST

PDB ID	:	5FJ5
EMDB ID:	:	EMD-3185
Title	:	Structure of the in vitro assembled bacteriophage phi6 polymerase complex
Authors	:	Ilca, S.; Kotecha, A.; Sun, X.; Poranen, M.P.; Stuart, D.I.; Huiskonen, J.T.
Deposited on	:	2015-10-06
Resolution	:	4.80 Å(reported)
Based on PDB ID	:	4K7H

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

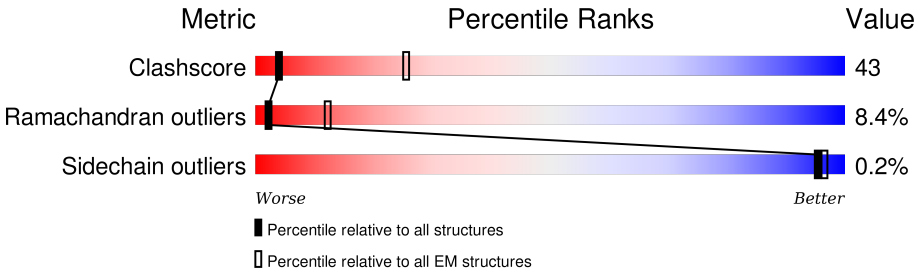
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	761	33% 59% 7% .
1	B	761	36% 59% 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 11840 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called MAJOR INNER PROTEIN P1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	761	Total	C	N	O	S	0	0
			5920	3741	1048	1109	22		
1	B	761	Total	C	N	O	S	0	0
			5920	3741	1048	1109	22		

There are 2 discrepancies between the modelled and reference sequences:

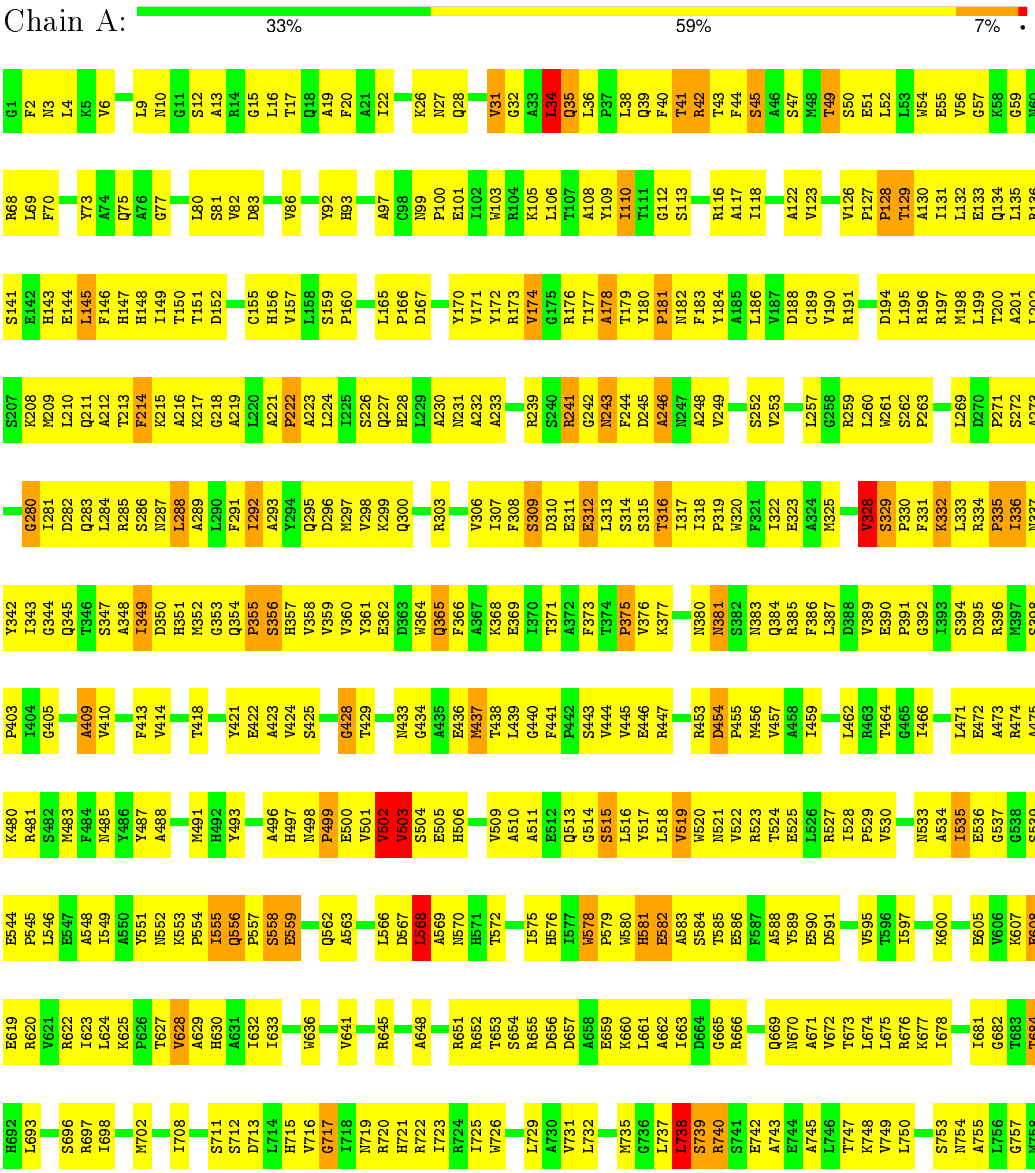
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P11126
B	1	GLY	-	EXPRESSION TAG	UNP P11126

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

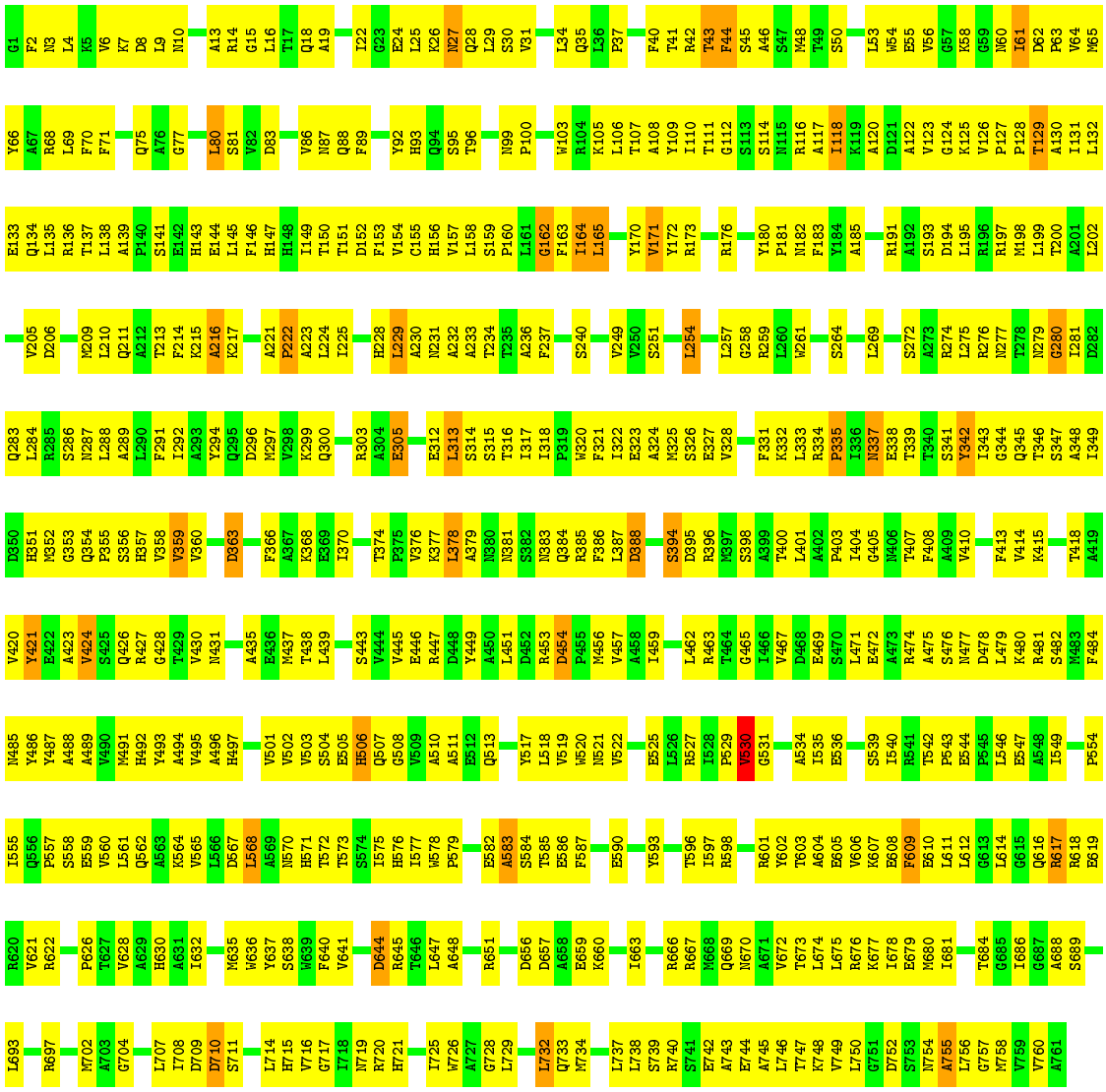
- Molecule 1: MAJOR INNER PROTEIN P1

Chain A:



- Molecule 1: MAJOR INNER PROTEIN P1

Chain B:


G1 F2 N3 L4 K5 V6 K7 D8 L9 N10 A13 R14 G15 L16 T17 Q18 A19 I22 G23 E24 L25 K26 N27 Q28 L29 S30 V31 L34 Q35 I36 P37 F40 T41 R42 T43 F44 S45 A46 S47 M48 T49 S50 L53 W54 E55 V56 G57 K58 G59 N60 I61 D62 P63 V64 W65
Y66 A67 R68 L69 F70 F71 Q75 A76 G77 L80 S81 V82 D83 V86 N87 Q88 F89 Y92 H93 Q94 S95 T96 N99 P100 W103 R104 K105 L106 T107 A108 Y109 I110 T111 G112 S113 S114 N115 R116 A117 I118 K119 A120 D121 A122 V123 G124 K125 V126 P127 P128 T129 A130 I131 L132
E133 Q134 L135 R136 T137 L138 A139 P140 S141 E142 H143 E144 L145 F146 H147 H148 I149 T150 T151 D152 F153 V154 C155 H156 V157 L158 S159 P160 L161 G162 F163 I164 L165 Y170 V171 Y172 R173 R176 Y180 P181 N182 F183 Y184 A185 R191 A192 S193 D194 L195 R196 R197 M198 L199 T200 A201 L202
V205 D206 M209 L210 Q211 A212 T213 F214 K215 A216 K217 A221 P222 A223 L224 I225 H228 L229 A230 N231 A232 A233 T234 T235 A236 F237 S240 V249 Y250 S251 L254 L257 G258 R259 L260 W261 S264 L269 S272 A273 R274 L275 R276 N277 T278 N279 G280 I281 D282
Q283 L284 R285 S286 N287 L288 A289 L290 F291 I292 A293 Y294 Q295 D296 M297 V298 K299 Q300 R303 A304 E305 E312 L313 S314 S315 T316 I317 I318 P319 W320 F321 I322 E323 A324 M325 S326 E327 V328 F331 K332 L333 R334 P335 I336 N337 E338 T339 T340 S341 Y342 I343 G344 Q345 T346 S347 A348 I349
D350 H351 M352 G353 Q354 P355 S356 H357 V358 V359 V360 D363 F366 A367 K368 E369 I370 T374 P375 V376 K377 L378 A379 N380 N381 S382 N383 Q384 R385 F386 L387 D388 S394 D395 R396 K397 S398 A399 T400 L401 A402 P403 I404 G405 N406 T407 F408 A409 V410 F413 V414 K415 T418 A419
V420 Y421 E422 A423 V424 S425 Q426 R427 G428 T429 V430 N431 A435 E436 M437 T438 L439 S443 V444 V445 E446 R447 D448 Y449 A450 L451 D452 R453 D454 P455 M456 V457 A458 I459 L462 R463 T464 G465 I466 V467 D468 E469 S470 L471 E472 A473 R474 A475 S476 N477 D478 L479 K480 R481 S482 N483 F484
N485 Y486 Y487 A488 A489 V490 M491 H492 Y493 A494 V495 A496 H497 V501 V502 V503 S504 E505 H506 Q507 G508 V509 A510 A511 E512 Q513 Y517 L518 V519 W520 N521 V522 E525 L526 R527 I528 P529 V530 G531 A534 I535 E536 S539 I540 R541 T542 P543 E544 P545 L546 E547 A548 I549 P554
I555 Q556 P557 S558 E559 V560 L561 Q562 A563 K564 V565 L566 D567 I568 A569 N570 H571 T572 T573 S574 I575 H576 I577 W578 P579 E582 A583 S584 T585 E586 F587 E590 Y593 T596 I597 R598 R601 Y602 T603 A604 B505 V606 K607 B508 F609 E610 L611 L612 G613 L614 G615 D616 R617 R618 E619
R620 V621 R622 P626 T627 V628 A629 H630 A631 I632 M635 W636 Y637 S638 W639 F640 V641 D644 R645 T646 L647 A648 R651 D656 D657 A658 E659 K660 I663 R666 R667 M668 Q669 N670 A671 V672 T673 L674 L675 R676 K677 I678 E679 M680 I681 T684 G685 I686 G687 A688 S689
L693 R697 M702 A703 G704 L707 I708 D709 D710 S711 L714 H715 V716 G717 I718 N719 R720 H721 I725 W726 A727 G728 L729 L732 Q733 M734 L737 L738 S739 R740 S741 E742 A743 E744 A745 L746 T747 K748 V749 L750 G751 D752 S753 N754 A755 L756 G757 M758 V759 V760 A761

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	16	Depositor
Minimum defocus (nm)	1100	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	160000	Depositor
Image detector	GATAN K2 (4K X 4K)	Depositor

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 2$	RMSZ	$\#\|Z\| > 2$
1	A	0.43	0/6040	0.70	7/8206 (0.1%)
1	B	0.40	0/6040	0.68	1/8206 (0.0%)
All	All	0.41	0/12080	0.69	8/16412 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	B	0	3
All	All	0	8

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	128	PRO	CA-C-N	6.74	132.03	117.20
1	A	128	PRO	C-N-CA	6.09	136.93	121.70
1	A	568	LEU	CA-CB-CG	5.92	128.90	115.30
1	A	502	VAL	C-N-CA	5.79	136.19	121.70
1	B	313	LEU	CA-CB-CG	5.67	128.35	115.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	328	VAL	Peptide
1	A	34	LEU	Peptide
1	A	35	GLN	Peptide
1	A	365	GLN	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	437	MET	Peptide

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5920	0	5913	515	0
1	B	5920	0	5913	504	0
All	All	11840	0	11826	1011	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 43.

The worst 5 of 1011 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:42:ARG:HD2	1:B:339:THR:H	1.30	0.96
1:A:117:ALA:H	1:A:221:ALA:H	1.10	0.93
1:B:117:ALA:HB1	1:B:221:ALA:HB1	1.50	0.91
1:B:80:LEU:HD12	1:B:81:SER:H	1.36	0.88
1:A:342:TYR:HA	1:A:559:GLU:HG3	1.57	0.86

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	759/761 (100%)	593 (78%)	92 (12%)	74 (10%)	1 14
1	B	759/761 (100%)	597 (79%)	108 (14%)	54 (7%)	1 22
All	All	1518/1522 (100%)	1190 (78%)	200 (13%)	128 (8%)	2 17

5 of 128 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	VAL
1	A	41	THR
1	A	129	THR
1	A	145	LEU
1	A	174	VAL

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	629/629 (100%)	627 (100%)	2 (0%)	94 96
1	B	629/629 (100%)	628 (100%)	1 (0%)	95 97
All	All	1258/1258 (100%)	1255 (100%)	3 (0%)	95 97

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	503	VAL
1	A	519	VAL
1	B	530	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	35	GLN
1	B	72	GLN
1	B	669	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	39	GLN
1	B	93	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.